CN 53-1189/P ISSN 1672-7673

基于日球物理学事件知识库的太阳活动识别*

李卫疆，亓 鑫，王 锋
(昆明理工大学云南省计算机技术应用重点实验室，云南 昆明 650500)

摘要：太阳图像中包含明显的太阳活动区域，以及无现象的太阳宁静区。从太阳图像中识别有效的太阳活动区域，是图像处理技术在天文研究中的典型应用。得益于美国太阳动力学天文台的日球物理学事件知识库提供的实时太阳观测数据，提出了一种基于日球物理学事件知识库的太阳活动识别方法。此方法获取 6 种太阳活动的信息(发生时间、位置、区域面积)，建立对应时间的全日面图像的尺度变换模型。结合位置与区域面积信息，对不同太阳活动进行梯度阈值分割，边界识别方法被用来定位和识别太阳活动的区域。然后，对太阳图像特征参数相关性的研究，得到每种太阳活动的最佳特征参数组合。方法实现了对太阳活动的精确定位和有效识别，为后续工作的开展提供了便利。此外，对不同太阳活动区域提取特定组合的特征，可以为基于内容的图像检索系统建立精简的图像特征集提供一种可行办法。

关键词：太阳活动定位；日球物理学事件知识库；梯度阈值；特征参数组合
中图分类号：TP391.41 **文献标识码**：A **文章编号**：1672-7673(2017)02-0251-10

基于内容的图像检索技术产生于 20 世纪 90 年代，它的出现解决了大规模数字图像检索的难题。随后，一大批研究性或商用的基于内容的图像检索系统(Content-Based Image Retrieval，CBIR)应运而生。区别于原有系统中对图像进行人工标注的做法，基于内容的检索技术自动提取每幅图像的视觉内容特征作为索引，如色彩、纹理、形状等。而针对大规模太阳活动的基于内容的图像检索系统也已出现，如太阳动力学天文台的基于内容的图像检索系统(SDO Content-Based Image Retrieval System)。

虽然现有的面向天文大数据的基于内容的图像系统能够满足在海量天文图像库中检索目标图像的基本需求，但是依然面临检索耗时、准确率不高的问题。出现这些问题的原因如下：(1)原始采集的太阳图像尺寸大，对应所提取的特征点明显增多，造成耗时增加；(2)太阳图像中呈现混淆难辨的太阳现象，所提取的特征没有针对性，造成匹配结果不理想。本文的研究对以上问题提供了解决思路。

本文的目标在于为建立太阳活动自动识别和检测系统探讨可行的办法，该系统的功能是自动识别一张给定的全日面太阳图像中所包含的太阳活动。建立这样一个系统所面临的难题：(1)不同太阳活动的物理特征差异很小，且同种太阳活动包含多种物理形态，使得太阳活动的判定不明确；(2)目前针对多种太阳活动识别的技术方法还不成熟；(3)特征库匹配时间开销大，检索效率低。

针对问题(1)，本文采用日球物理学事件知识库提供的实时太阳观测数据，数据由太阳动力学天文台的特征检测小组(Feature Finding Team，FFT)模块自动获取，避免了人为干预，数据包括太阳实时活动的发生时间、在太阳表面的发生位置以及活动类型；针对问题(2)，本文采用网格结构划分图像，使用梯度阈值法分割太阳活动的目标区域，使用边界识别方法最小外接矩形(Minimum Bounding Rectangle，MBR)和八连通链码获取目标区域的轮廓特征数据；针对问题(3)，本文研究了 6 种图像参数(主要是纹理特征)所属特性，探讨图像参数与太阳活动的相关性，找出每种太阳活动对应的最优参数或参数组合，并依次对包含 6 类太阳活动的目标区域图像集进行特征提取，用于建立特征库。

---

* 基金项目：国家自然科学基金 (61363045) 资助.
收稿日期：2016-07-04；修订日期：2016-07-28
作者简介：李卫疆，男，副教授. 研究方向：信息检索. Email：aininala1@ sina.cn

1 相关工作

太阳动力学天文台[1]是美国国家航空航天局一个为期 10 年的太阳观测任务。太阳动力学天文台的科学目标是在小尺度的时间和空间下以多波段研究太阳大气层，从而了解太阳对地球和近地球太空区域的影响。大气成像组件(Atmospheric Imaging Assembly，AIA)用于拍摄高时间与高空间分辨率的完整太阳影像。事件检测系统(Event Detection System，EDS)利用特征检测小组自动产生特征数据，并传送到日球物理学事件知识库(Heliophysics Event Knowledgebase，HEK)。文[2]利用分层迭代法以及最优化方法对日球物理学事件知识库提供的太阳观测数据进行太阳活动间的相关性研究，最终求解网络流。目前，该方法仅对太阳活动区和日冕洞两种太阳活动的追踪效果较好。文[3]结合日球物理学事件知识库的太阳观测数据，对方形活动区域提取 10 种特征参数，构造用于太阳活动区域检索的特征描述符，并研究了对应太阳活动的最佳相似匹配方法。结果显示，对于除太阳活动区、耀斑和 S 型结构 3 种太阳活动外，该方法还不能获取满意的检索准确率。

针对单一太阳活动现象的自动识别成为非常流行的研究课题。文[4]利用神经网络贝叶斯(Bayes)推理和形状关系自动识别包括太阳黑子、耀斑和日冕物质抛射。文[5]使用局部阈值和区域生长法在全日面图像中识别太阳活动区。文[6]训练了神经网络识别太阳图像的隐藏背景中的暗条。文[7]提出了一种自动的暗条检测方法，这种方法超越了典型的暗条检测，它确定暗条的脊和方向角，发现暗条手性。文[8]提出了一种自动识别耀斑算法，这种算法在噪声和背景有亮线影响的情况下效果很好。对于日冕物质抛射的识别，文[9]提出了一种基于数据立方体的算法，并利用小波变换分析识别日震。目前的方法还仅限于识别单一太阳活动或物理特征相近的几种太阳活动，适用性不佳，无法满足海量天文图像处理的需求。

虽然，目前的识别算法结合了太阳活动的物理特征，并综合运用了图像处理的各种技术，因此对单一太阳活动识别取得了比较好的结果，但是还没有通行的方法针对全日面图像上的多种太阳活动进行识别。随着天文图像数量的激增，现有的太阳活动识别方法在时效性和识别精度方面不能满足多波段太阳图像的识别任务。而日球物理学事件知识库实时提供太阳活动的精准数据源，这些数据源记录了太阳活动物理特征的重要参数，可以为相关研究工作提供有力的参考[10]。现有的方法中结合日球物理学事件知识库与图像处理方法用于太阳活动定位和识别的相关研究还不多。本文介绍一种基于日球物理学事件知识库的太阳活动定位和识别的方法，以实现对全日面图像太阳活动进行自动识别。

本文从日球物理学事件知识库获取 6 类太阳活动的实时观测数据，包括活动区(Active Region，AR)、日冕洞(Coronal Hole，CH)、浮现区(Emerging Flux，EF)、耀斑(Flare，FL)、黑子(Sunspot，SS)、S 型结构(Sigmoid，SG)，相比其它类型的太阳活动，以上 6 类太阳活动发生频繁，且持续时间长。

太阳动力学天文台位于太空中的成像系统大气成像组件获取不同波段的紫外线光谱成像，拍摄的太阳图像不受大气透明度和大气扰动等因素的影响，但是太阳全日面上的亮度在紫外线光谱通道边缘变亮，出现临边增亮现象(Limb Brightening)。本文首先对太阳图像进行标准化处理，去除太阳边缘亮圈，避免非日面区域对后续工作的影响。然后，从日球物理学事件知识库采集 365 张全日面图像的所有太阳活动的相关信息，包括太阳活动的发生时间、发生位置、覆盖面积等，为了将这些信息应用于标准化后的全日面图像，需要建立对应时间的全日面图像的尺度变换模型，产生的新位置信息和面积信息将应用于本文的图像分割方法中。

采用梯度阈值法，结合不同太阳活动的物理特性与尺度变化后的新数据，对不同太阳活动分区域进行分割，将分割后的太阳活动目标区域以两种边界识别方法标注显示，记录目标区域的轮廓信息。

为了建立用于基于内容的图像检索系统的精简图像特征集，本文研究了太阳图像 6 种特征的相关性。对 2015 全年的全日面太阳图像提取 6 种特征参数，通过分析 6 种特征相关性，找出相关性最强的图像特征组合后就能实现降低特征参数的维度。结合目标区域的轮廓信息，提取目标区域内的特征组合，建立特征集。

本文提出一种基于日球物理学事件知识库的太阳活动识别方法，该方法结合日球物理学事件知识库提供的太阳活动观测数据，利用多种图像处理方法，实现了对尺度可变换的全日面图像中多种太阳活动的准确定位与识别。实验结果表明，该方法在太阳活动识别准确度及时间开销方面取得满意的效果。

2　太阳活动识别方法

2.1　Hough 变换检测圆心和半径

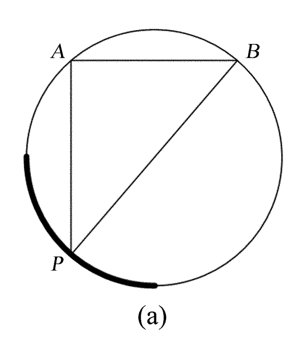
Hough 变换是模式识别领域对二值图像进行圆检测的有效方法，本文提出一种利用圆斜率特性的改进 Hough 变换方法，用于检测太阳圆心和半径以及去除边缘亮圈。基本思想是将图像的空间域变换到参数空间，用大多数边界点满足的某种参数形式描述图像中的区域边界。通过设置累加器进行累加，求得峰值对应的点就是所需要的信息。

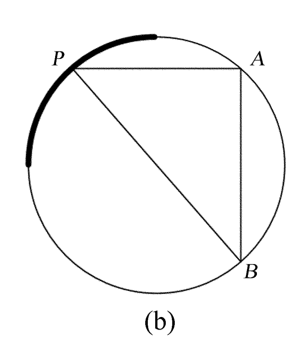
2.1.1　圆的几何特性分析

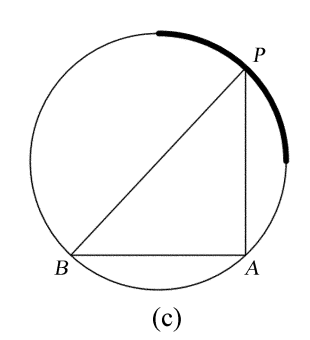
圆周上存在 4 个点 $P(x_{\min}, y_1)$、$P(x_u, y_{\min})$、$P(x_{\max}, y_r)$、$P(x_d, y_{\max})$ 分别对应圆的左端、上端、右端、下端的极值点，除这 4 个极值坐标点外，对于圆周上的任意一点 $P(x, y)$，过点 P 存在一条垂直方向的线段 L_A 与一条水平方向的线段 L_B，L_A、L_B 与圆的交点分别是 A、B，ΔPAB 是过点 P 的内接直角三角形，根据圆的几何特性：圆的内接直角三角形的斜边过圆心，因此，AB 线段的中点就是圆心点。限定 P 的移动范围，构造 4 种圆内接直角三角形(如图 1)：

(a)为 $\{P \mid x_{\min} < x < x_d,\ y_1 < y < y_{\max}\}$；(b)为 $\{P \mid x_{\min} < x < x_u,\ y_{\min} < y < y_1\}$；

(c)为 $\{P \mid x_u < x < x_{\max},\ y_{\min} < y < y_r\}$；(d)为 $\{P \mid x_d < x < x_{\max},\ y_r < y < y_{\max}\}$。







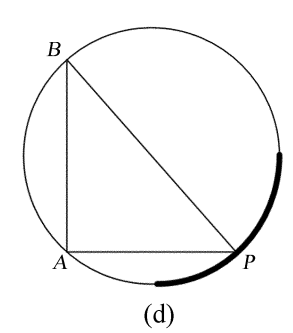


图 1　4 种圆内接直角三角形

Fig. 1　Four kinds of circle inscribed right-triangle

设 $B(x_1, y_1)$，则 $\overline{PB}$ 的中点坐标为 $x_0 = (x + x_1)/2$，$y_0 = (y + y_1)/2$。

2.1.2　Hough 变换检测圆心和半径

(1)利用索贝尔(SOBEL)算子从原始图像中检测出边缘信息，并转化边界为单一像素的边缘图，然后将图像二值化(即设定边缘像素值为 1，其他像素值为 0)，将边缘像素坐标存入数组 $V(x, y)$。

(2)设置 4 个累加器数组 $C(x, y)$(初始值为 0)，按照 2.1.1 节方法，逐个判断 $V(x, y)$ 中 P 的所属情况，计算中点坐标(x_0, y_0)，并将对应情况的累加器数组 $C(x, y)$ 加 1。累加器数组 C 中的最大值对应的坐标即为圆心(p, q)。

(3)将圆心坐标(p, q)和 $V(x, y)$代入圆方程，利用 Hough 变换对参数 R 进行统计，当 R 的统计值大于 80%(即 $V(x, y)$ 中 80%的点落在圆周上)，R 即为圆的半径。

(4)根据圆心和半径，将图像超出半径的区域设置为背景，实现去除边缘亮圈的目的。

2.2　图像建模方法

2.2.1　尺度变换

对 365 张原始全日面图像进行标准化处理后，图像的大小尺度发生变化，除太阳活动的发生时间

T 不变外，太阳活动的发生位置、覆盖面积随图像尺度大小的变化而变化。设原始采集信息中太阳活动的发生位置为(x_pre，y_pre)，覆盖面积为 $Area_pre$，图像中太阳半径为 R_pre，经 2.1 节标准化处理后，设太阳活动的新位置为(x_new，y_new)，覆盖面积为 $Area_new$，图像中太阳半径为 R_new。变换过程如下：

如图 2(a)，假设太阳活动的发生位置位于太阳边缘点，那么存在以下关系：

$$\begin{cases} x_new = x_pre\ (R_new/R_pre) \\ y_new = y_pre\ (R_new/R_pre) \end{cases} . \tag{1}$$

如图 2(b)，假设太阳活动为阴影部分，其面积 S 的计算如下：

$$S = \int_{\theta}^{\frac{\pi}{2}} 2\pi rR\mathrm{d}\theta = \int_{\theta}^{\frac{\pi}{2}} 2\pi R^2\cos\theta\mathrm{d}\theta = 2\pi R^2\int_{\theta}^{\frac{\pi}{2}}\cos\theta\mathrm{d}\theta = 2\pi R^2(1 - \sin\theta) = 2\pi RH . \tag{2}$$

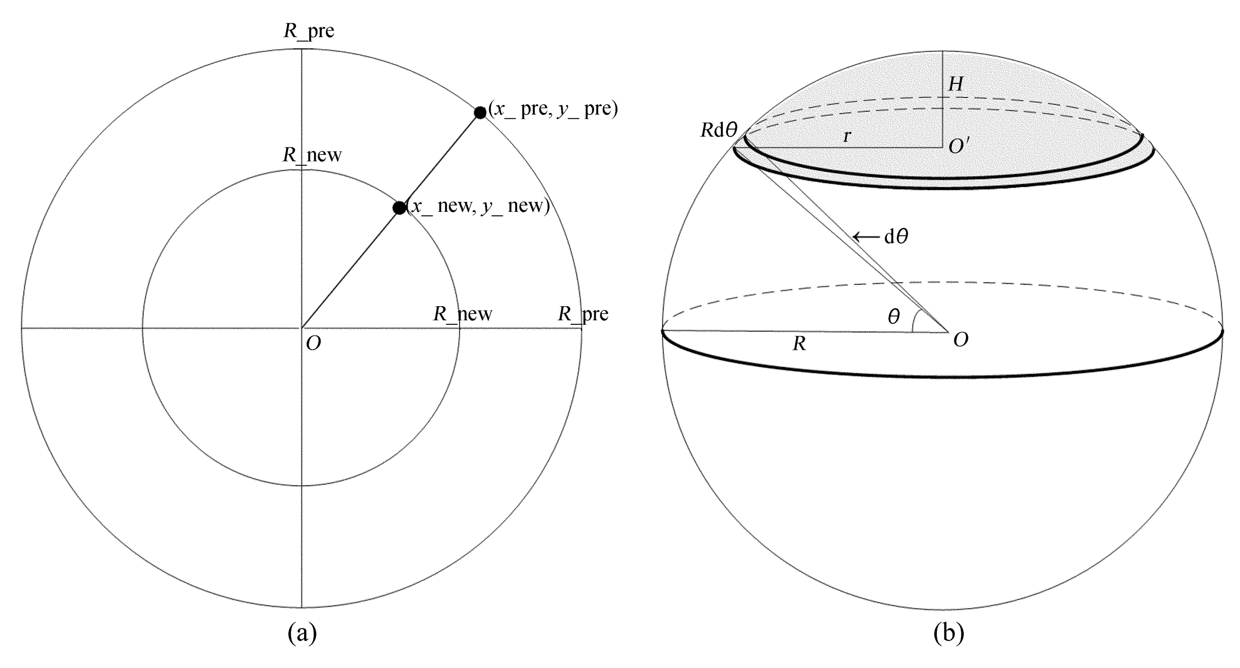


图 2　尺度变换的几何模型

Fig. 2　Geometric model for scale transformation

由(2)式推导太阳活动的覆盖面积：

$$\begin{cases} Area_pre = 2\pi R_preH \\ Area_new = 2\pi R_newh \end{cases} \tag{3}$$

令 $\kappa = \dfrac{R_pre}{R_new} = \dfrac{H}{h}$，进而推导出如下关系：

$$\frac{Area_pre}{Area_new} = \kappa^2 . \tag{4}$$

由 2.1 节得到尺度变化后的新图，建立用于图像分割的直角坐标系(图像的首个像素点作为坐标系原点)，坐标系变化后，对太阳活动的位置坐标(x，y)作再次调整：

$$\begin{cases} x_new = R_new + x_new \\ y_new = R_new - y_new \end{cases} \tag{5}$$

2.2.2　网格划分

网格划分是常规的图像处理方法，对标准化处理后的图像用 4 × 4 像素的网格进行划分，考虑到低维的数据结构要比高纬的数据结构在运算效率方面表现更优，本文采用之前研究中的处理方法，对网格数据进行降维操作和存储[11]。

2.3　边界识别方法

最小外接矩形和八连通链码是图像分类和识别研究常用的方法，被用来提取目标区域的几何特征[12]。本文最小外接矩形作为修正和划分梯度阈值的工具，与链码共同识别 6 种太阳活动区域。特别地，最小外接矩形信息存储在 $POINT(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4, x_1, y_1)$ 中，而链码信息采用 $POLYGON(x_1, y_1, \omega_1, x_2, y_2, \omega_2, \dots, x_n, y_n, \omega_n)$ 压缩存储，其中 (x, y, ω) 表示网格首像素的坐标及方向。

2.4　梯度阈值分割法

本文采用梯度阈值分割法，结合 6 种太阳活动的物理特性，在满足判定条件的情况下，分配不同的阈值。根据 2.2.1 节新坐标系中太阳活动位置坐标(x_new, y_new)，将该点的灰度值作为初始阈值，对图像的网格单元进行初步分割，利用网格单元的最小外接矩形修正阈值。梯度阈值分割法过程如下：

输入：一张经过 Hough 变换及尺度变换处理后的图像 IMG

输出：一张所有太阳活动区域被标注后的图像

步骤：

1：Convert IMG to a gray image
2：IMG is divided into M rows and N columns of cells
3：IMG→Block(pixel_row, pixel_col, cell_row, cell_col)
　　→Block(pixel_row, pixel_col, index)
4：Block(pixel_row, pixel_col, index)→img
5：Set img(x_new, y_new) to the initial threshold T_0
6：　if ~exist MBR(img)
　　　set threshold to $T_1(T_1 < T_0)$
7：　else if exist MBR(img) ∧ Area(MBR) / Area_new>threshold_area
　　　set threshold to $T_2(T_2 > T_1)$
8：　else exist MBR(img) ∧ Area(MBR) / Area_new<threshold_area
　　　set threshold to $T_3(T_1 < T_3 < T_2)$
9：count the number: [Block(pixel_row, pixel_col, index)>T] ($T \in [T_0, T_1, T_2, T_3]$)
10：　if T > threshold_number
11：　　set Block(pixel_row, pixel_col, index) to 255
12：　else set Block(pixel_row, pixel_col, index) to 0
13：　　count the density of connected square grids
14：　if density < threshold_density
15：　　remove the grid area
16：　else end

经过 Hough 变换、尺度变换等处理后，根据太阳活动的位置信息，将新的图像分区域进行网格划分，并依次对 6 种太阳活动进行梯度阈值分割，使用最小外接矩形和八连通链码标注太阳活动区域。其中，步骤 5 将太阳活动的位置点的灰度值作为初始阈值 T_0，步骤 6 设置判定条件(1)：若初次分割后的识别区域不存在最小外接矩形，设定新阈值 $T_1(T_1 < T_0)$。步骤 7 设置判定条件(2)：若初次分割后的识别区域存在最小外接矩形，并且 *Area*(MBR)与 *Area*_new 面积的比值大于经验阈值 threshold_area，设定新阈值 $T_2(T_2 > T_1)$。步骤 8 设置判定条件(3)：若初次分割后的识别区域存在最小外接矩形，并且 *Area*(MBR)与 *Area*_new 面积的比值小于经验阈值 threshold_area，设定新阈值 $T_3(T_1 < T_3 < T_2)$。步骤 6~8 根据 3 种判定条件，设置梯度阈值($T_1 < T_0 < T_3 < T_2$)。对于每种太阳活动，以上 3 种判定条件可灵活调整，以及分配不同的梯度阈值。本文中，threshold_number 取经验阈值 0.15，threshold_density 取经验阈值 0.5，threshold_area⊆[0.3, 0.7]。不同太阳活动的物理特性存在差异，太阳活动区、浮现区、耀斑以及 S 型结构相较于黑子、日冕洞，亮度更加明显(耀斑是最显著的太阳活动，黑子、日冕洞亮度则很低)。因此，选取合适的阈值对太阳图像分割产生有益的效果。

2.5 特征参数相关性评价方法

首先介绍 6 种图像特征参数。表 1 中，L 为太阳活动区域的像素个数，z_i 表示第 i 个像素的灰度值，$p(z_i)$ 表示灰度 z_i 在灰度直方图中的比值。

对太阳图像自动特征提取方法已经有了部分研究[13-14]，这些方法从多种固定参数组合中筛选表现优异的参数组合，而涉及各特征参数间关联性的研究则较少。为了研究各特征参数间的关联性，本文采用如下方法。

(1)提取太阳图像活动区域 X(外接矩形的标注范围)的 6 种特征参数，$q_i(X_k)$ 表示提取第 k 个活动区域的第 i 种参数。

(2)计算太阳活动区域 X 中第 i 个参数和第 j 个参数的协方差矩阵：

表 1 图像特征参数列表

Table 1 List of Extracted Image Parameters

编号	名称	公式
P1	熵	$E=-\sum_{i=1}^{L}p(z_i)\log_2 p(z_i)$
P2	均值	$m=\frac{1}{L}\sum_{i=1}^{L}z_i$
P3	标准差	$\sigma=\sqrt{\frac{1}{L}\sum_{i=1}^{L}(z_i-m)^2}$
P4	偏度系数	$u_3=\sum_{i=1}^{L}(z_i-m)^3p(z_i)$
P5	均匀性	$U=\sum_{i=1}^{L}p^2(z_i)$
P6	相对平滑度	$R=1-\frac{1}{1+\sigma^2(z)}$

$$\sum\nolimits_{ij}=\frac{1}{DN}\sum_{n=1}^{N}\sum_{k=1}^{D_n}[q_i(X_k)-u_i]\cdot[q_j(X_k)-u_j]\,, \tag{6}$$

其中，$u_i=\frac{1}{DN}\sum_{n=1}^{N}\sum_{k=1}^{D_n}q_i(X_k)$；$N=365$；$D_n$ 为第 n 张太阳图像所包含的活动数量。

(3)计算各参数间的相关矩阵 R_{ij}，反映了第 i 和第 j 两个参数间的相关程度，即 R_{ij}越大，两特征越相关。

$$R_{ij}=\frac{\sum_{ij}}{\sqrt{\sum_{ii}\sum_{jj}}}\,. \tag{7}$$

3 实 验

3.1 数据集

本文图像数据集来源于太阳动力学天文台的大气成像组件(2015 年)①，共 365 张太阳全日面图像，图像基本信息如表 2。

表 2 太阳活动的基本信息

Table 2 A summary of event types

太阳活动	来源	报道次数	采集模块
太阳活动区	193°A	20 226	空间可能性聚类算法模块
日冕洞	193°A	16 590	空间可能性聚类算法模块
耀斑	131°A	21 774	耀斑检测触发模块
S 型结构	131°A	8 538	S 型结构检测模块
浮现区	日震与磁成像仪	10 467	浮现区检测模块
黑子	日震与磁成像仪	4 493	欧洲太阳黑子观测模块

太阳动力学天文台通过仪器捕捉不同波段的太阳紫外线，由大气成像组件进行光谱成像，合成具有明显太阳活动的图像，不同太阳活动被事件检测系统的相应模块报道[15]。例如，太阳黑子由日震与磁成像仪(Helioseismic and Magnetic Imager，HMI)检测，这种现象在 2015 年全年被 EGSO_SFC 模块报道了 4 493 次。

① http://sdowww.lmsal.com/suntoday_v2/

3.2 太阳活动区域定位与识别

本文采用 3 种波段紫外线(波长 211、193、171)叠加后形成的图像，这些图像中太阳活动区、耀斑、日冕洞、浮现区 4 种太阳活动的物理特征较为清晰。利用日球物理学事件知识库提供的信息，结合多种方法，在全日面图像上完成了太阳活动的定位和识别工作。最小外接矩形准确定位了完整的太阳活动区域，八连通链码较好地展示了太阳活动的形态特征。如图 3。

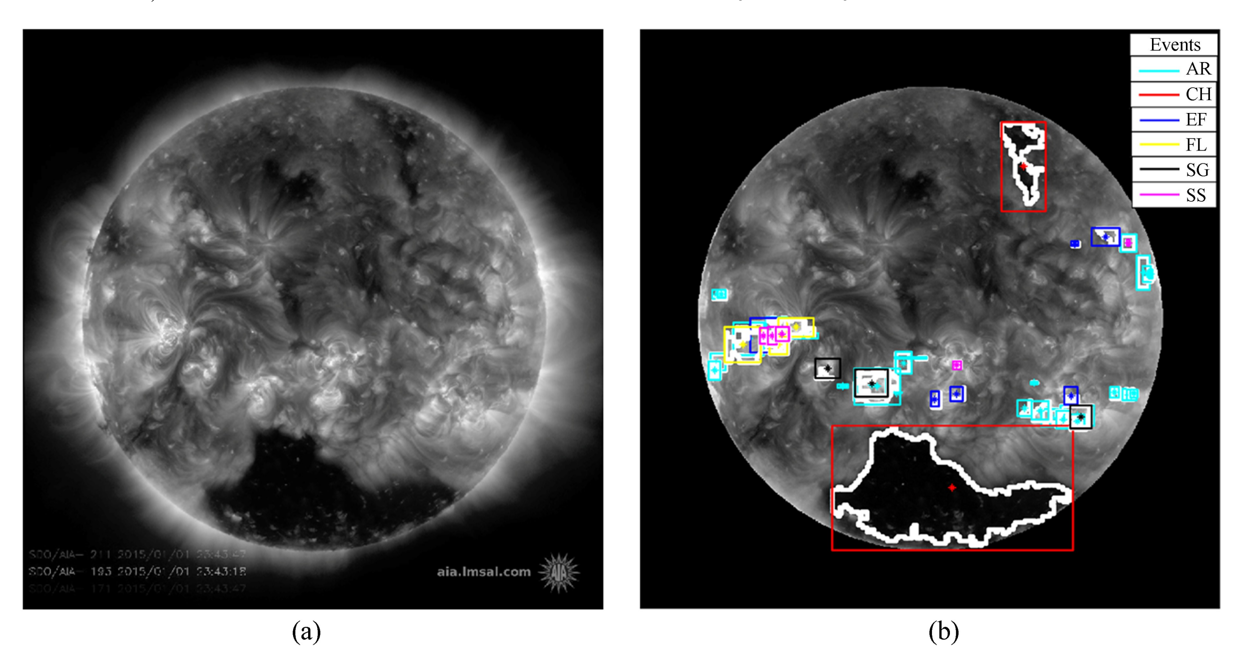


图 3 (a) 来源于大气成像组件的全日面图像；(b) 太阳活动分割与识别

Fig. 3 (a) Original image of the Sun taken on 1 January 2015 from AIA; (b) Segmentation and Recognition of solar activities

3.3 识别准确度评估

本文进一步统计了 2015 年 1 月 31 天的 6 种太阳活动的分割比例(分割比例是指八连通链码识别区域面积与最小外接矩形定位区域面积的比值)，用于评估太阳活动识别的准确度。具体过程：采用 3 种方法(种子区域生长法(SRG)、分水岭分割法(WaterShed)以及梯度阈值分割法(GTS))对太阳活动定位区域进行活动识别，然后，分别统计 3 种方法应用于图像集后的分割比例。结果见图 4 和图 5。

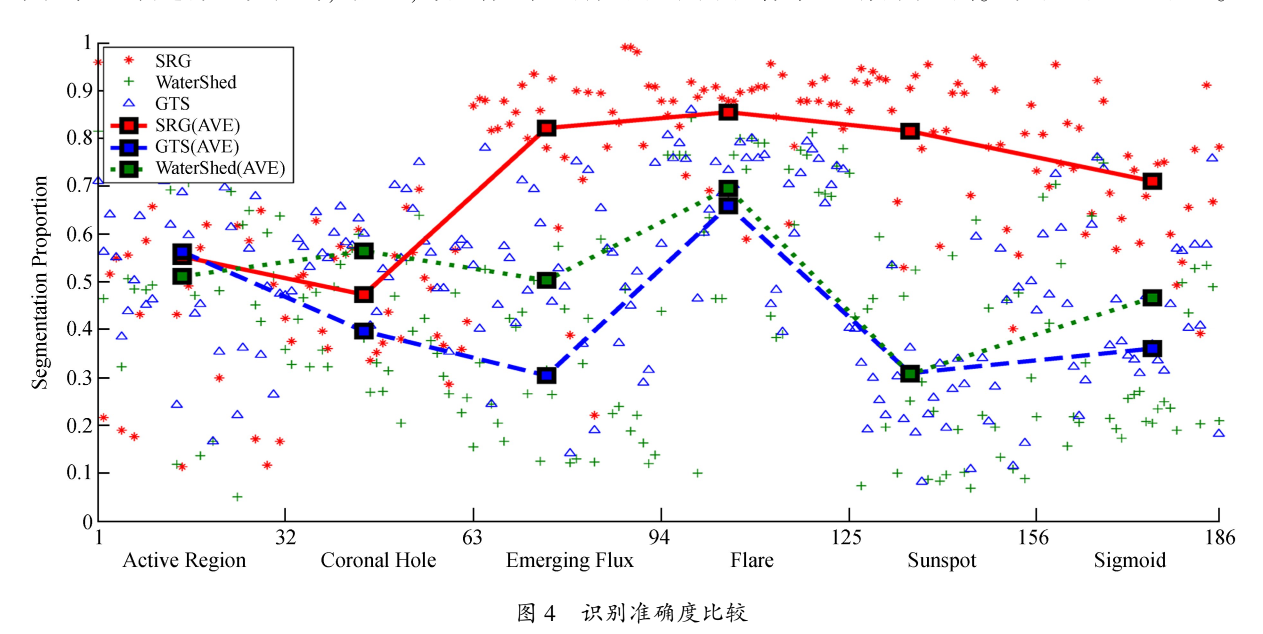


图 4 识别准确度比较

Fig. 4 Comparison of recognition accuracy of three methods

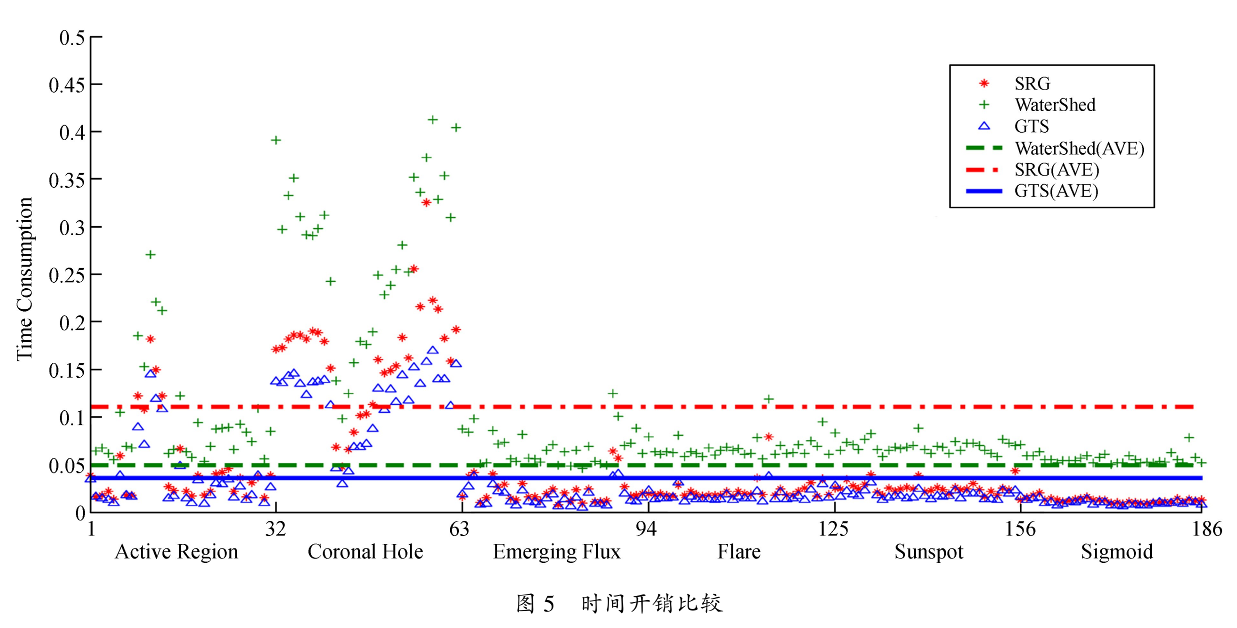


图 5 时间开销比较

Fig. 5 Comparison of time consumption of three methods

图 4 对比了 3 种方法的分割比例，每种分割方法应用于 186 个太阳活动区域，对离散数据点的进一步分析发现，使用种子区域生长法得到的 186 组数据中，共有 70 组分割比例在 0. 3～0. 7 范围内，其它 116 组数据是由于图像过度分割或者图像空洞等原因产生的，这些数据被认为无法满足识别要求，采用种子区域生长法的识别准确率为 37. 6%。相应地，分水岭分割法的识别准确率为 48. 9%(共有 91 组数据符合要求)，梯度阈值分割法的识别准确率最高，为 65. 6%(共有 122 组数据符合要求)。对于分割比例过低或者过高的情况，可通过 2. 4 节调整梯度阈值改善结果。

图 5 比较了 3 种分割方法的时间开销，其中值得注意的是，太阳活动区以及日冕洞这两种太阳活动的发生范围较大，因而在时间开销上大于其余 6 种太阳活动。统计结果显示，种子区域生长法的平均时间开销为 0. 111 s，分水岭分割法和梯度阈值分割法的平均时间开销分别为 0. 049 s、0. 035 s。

3. 4 特征参数相关性评价

由 2. 5 节方法对 2015 年 1 月 31 天的太阳图像进行特征相关性计算，将 6 种太阳活动图像各特征参数相关性结果用热度图表示(白色为 1，黑色为 0)，如图 6。

将 6 种太阳活动不同特征间相关性进行可视化表征，得到以下结果：太阳活动区图像的可选特征参数组合包括(P1，P4)、(P2，P3)、(P4，P5，P6)；日冕洞图像的可选特征参数组合包括(P1，P5，P6)、(P4，P5)；浮现区图像的可选特征参数组合包括(P1，P3)、(P2，P4，P6)；耀斑图像的可选特征参数组合包括(P1，P2)、(P3，P4，P6)；黑子图像的可选特征参数组合包括(P4，P5，P6)、(P3，P4)；S 型结构图像的可选特征参数组合包括 (P1，P3，P5)、(P4，P5)。对太阳活动区域提取特征参数，应优先选择耗时最少的组合。

4 总 结

本文从太阳动力学天文台的日球物理学事件知识库采集 6 种太阳活动的基本信息，通过对图像建立尺度变换模型，结合每种太阳活动的物理特征对图像活动区域进行梯度阈值分割，使用最小外接矩形和弗里曼(Freeman)码实现对太阳活动区域的定位与识别。对太阳图像各特征参数的相关性进行研究，得出了合适的参数组合。研究结果说明基于日球物理学事件知识库的太阳活动识别方法的可行性。但仍有可改进的地方，主要是对各太阳活动物理特征在灰度空间的表现，需要加深了解，有助于改进梯度阈值分割方法，进一步提高对太阳活动识别的准确度。未来工作将对太阳活动区域提取特定的组合参数，应用精简数据集建立基于内容的图像检索系统，并测试在检索耗时与检索准确率等方面的表现。

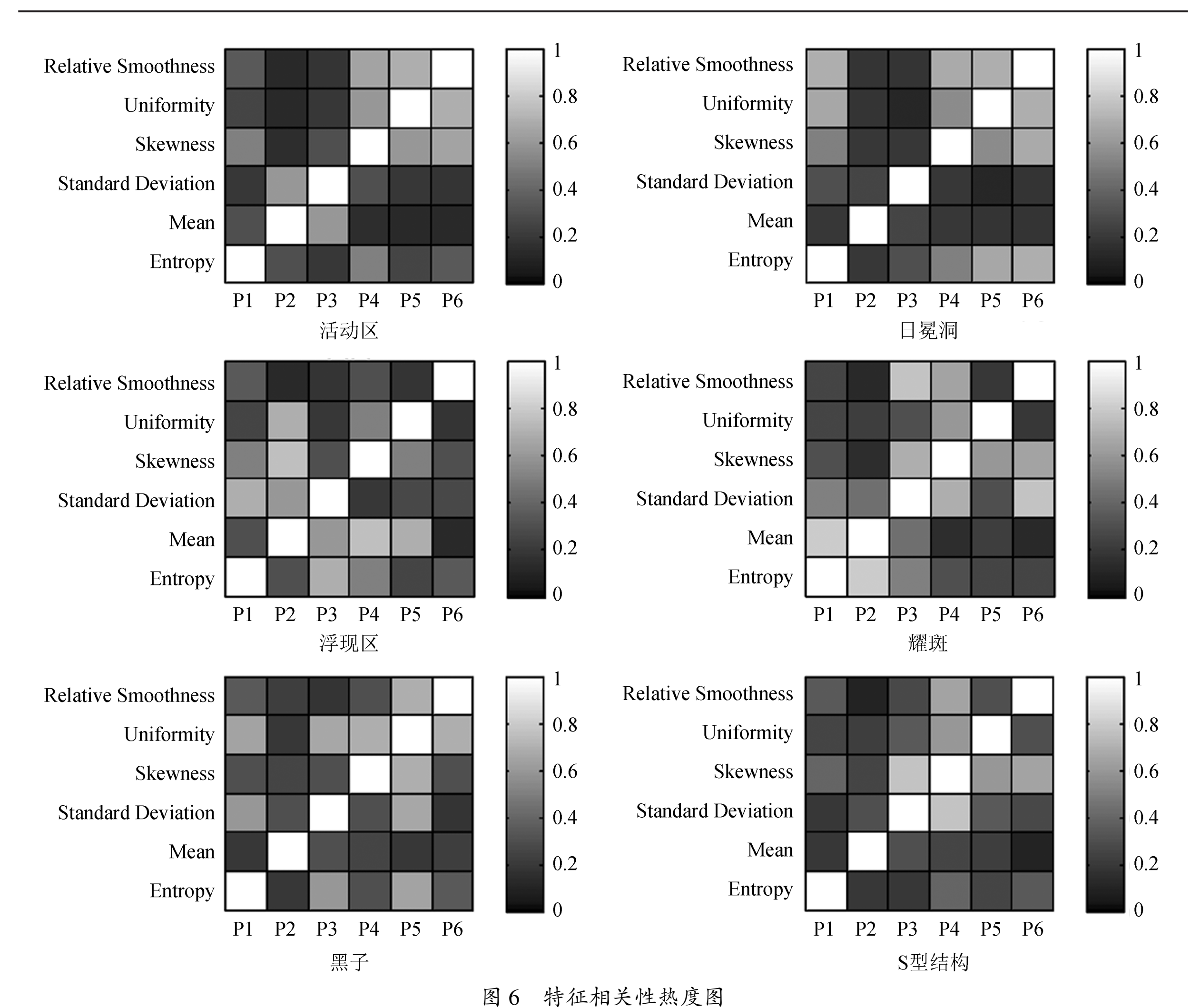


图 6　特征相关性热度图

Fig. 6　The heat map of features correlation

参考文献:

[1]　Pesnell W D, Thompson B J, Chamberlin P C. The Solar Dynamics Observatory (SDO) [J]. Solar Physics, 2012, 275(1-2): 3-15.

[2]　Kempton D, Pillai K G, Angryk R. Iterative refinement of multiple targets tracking of solar events [C] // 2nd IEEE International Conference on Big Data. 2014.

[3]　Banda J M, Liu C, Angryk R A. Region-based querying of solar data using descriptor signatures [C] // IEEE Conference on Data Mining Workshops. 2013.

[4]　Zharkova V V, Ipson S S, Benkhalil A, et al. Feature recognition in solar images [J]. Artificial Intelligence Review, 2005, 23(3): 209-266.

[5]　Benkhalil A, Zharkova V V, Zharkov S, et al. Automatic identification of Active Regions (Plages) in the full-disk solar images using local thresholding and region growing techniques [C] // Proceedings of the AISB'03 Symposium on Biologically-inspired Machine Vision, Theory and Application. 2003.

[6]　Zharkova V V, Schetinin V. Filament recognition in solar images with the neural network technique [J]. Solar Physics. 2005, 228(1-2): 137-148.

[7]　Bernasconi P N, Rust D M, Hakim D. Advanced automated solar filament detection and characterization code: description, performance, and results [J]. Solar Physics, 2005, 228(1): 97-117.

[8] Christe S, Hannah I G, Krucker S, et al. RHESSI microflare statistics. I. flare-finding and frequency distributions [J]. The Astrophysical Journal, 2008, 677(2): 1385-1394.

[9] Savcheva A, Cirtain J, Deluca E E, et al. A study of polar jet parameters based on Hinode XRT observations [J]. Publications of the Astronomical Society of Japan, 2007, 59(SP3): S771-S778.

[10] Hurlburt N, Cheung M, Schrijver C, et al. Heliophysics Event Knowledgebase for the Solar Dynamics Observatory (SDO) and beyond [J]. Solar Physics, 2012, 275(1-2): 67-78.

[11] 李卫疆, 亓鑫, 王锋. 基于方形网格结构的太阳活动目标检测方法 [J]. 天文研究与技术, 2016, 13(4): 481-488.
Li Weijiang, Qi Xin, Wang Feng. A square grid structure for target detection of solar activities [J]. Astronomical Research & Technology, 2016, 13(4): 481-488.

[12] 卢蓉, 范勇, 陈念年, 等. 一种提取目标图像最小外接矩形的快速算法 [J]. 计算机工程, 2010, 36(21): 178-180.
Lu Rong, Fan Yong, Chen Niannian, et al. Fast algorithm for extraction Minimum Enclosing Rectangle of target image [J]. Computer Engineering, 2010, 36(21): 178-180.

[13] Vasconcelos N, Vasconcelos M. Scalable discriminant feature selection for image retrieval and recognition [C]// IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2004.

[14] Deselaers T, Keysers D, Ney H. Features for image retrieval: an experimental comparison [J]. Information Retrieval, 2008, 11(2): 77-107.

[15] Martens P C H, Attrill G D R, Davey A R, et al. Computer vision for the Solar Dynamics Observatory (SDO) [J]. Solar Physics, 2012, 275(1-2): 79-113.

Recognition of Solar Activities Based on Heliophysics Event Knowledgebase

Li Weijiang, Qi Xin, Wang Feng

(Key Laboratory of Applications of Computer Technologies of the Yunnan Province, University of Science and Technology of Kunming, Kunming 650500, China, Email: aininala1@sina.cn)

Abstract: Solar image conveys obvious active regions and peaceful regions of no activity. Recognizing useful solar activities from solar image is a typical application of image processing in the astronomical research scope. Benefiting from real-time observation data provided by Heliophysics Event Knowledgebase (HEK) of Solar Dynamics Observatory (SDO), this paper proposes a recognition method of solar activities based on HEK. In this method, we gather six kinds of solar activities datas (date, location, area), then we design a scaling transform model for full-disk solar image of corresponding date. Combining location data and area data, we design different gradient thresholds to segment different solar activities regions. Two kinds of boundary identification method are used to locate and recognize solar activities regions respectively. Furthermore, we study the correlation of six kinds of solar image features, and give combinations of features for every solar activity. The method in the paper realizes the precise positioning and efficient recognition, which lies the foundation for further work. Besides, extracting particular combination of features from different solar activities regions also provides a practical approach of constructing a condensed feature set for CBIR system.

Key words: Solar activity location; HEK; Gradient thresholds; Combination of features